

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/S25,674
Source: PCT
Date Processed by STIC: 3-8-05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 03/08/2005

PATENT APPLICATION: US/10/525,674

TIME: 08:02:24

Input Set : A:\Final Sequence List - 13111-2-US.txt

Output Set: N:\CRF4\03082005\J525674.raw

3 <110> APPLICANT: Kroger, Burkhard
 4 Zelder, Oskar
 5 Kolpprogge, Corinna
 6 Schroder, Hartwig
 7 Hafner, Stefan
 9 <120> TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
 10 Sulphur (Meta)
 12 <130> FILE REFERENCE: 13111-00002-US
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/525,674
 C--> 14 <141> CURRENT FILING DATE: 2005-02-24
 14 <150> PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
 15 <151> PRIOR FILING DATE: 2003-08-26
 17 <150> PRIOR APPLICATION NUMBER: DE 102 39 073.8
 18 <151> PRIOR FILING DATE: 2002-08-26
 20 <160> NUMBER OF SEQ ID NOS: 69
 22 <170> SOFTWARE: PatentIn version 3.3
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 1104
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Corynebacterium diphtheriae
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (1)..(1101)
 34 <223> OTHER INFORMATION: RDI00386
 36 <400> SEQUENCE: 1
 37 atg ctc acc acc aca ggg acg ctc acg cac caa aaa atc gga gac ttt 48
 38 Met Leu Thr Thr Thr Gly Thr Leu Thr His Gln Lys Ile Gly Asp Phe
 39 1 5 10 15
 41 tac acc gaa gcc gga gcg acg ctt cac gac gta acc atc gcc tac caa 96
 42 Tyr Thr Glu Ala Gly Ala Thr Leu His Asp Val Thr Ile Ala Tyr Gln
 43 20 25 30
 45 gca tgg ggc cac tac acc ggc acc aat ctc atc gtt ctc gaa cat gcc 144
 46 Ala Trp Gly His Tyr Thr Gly Thr Asn Leu Ile Val Leu Glu His Ala
 47 35 40 45
 49 ctg acc ggc gac tct aac gct att tca tgg tgg gac gga ctg att ggc 192
 50 Leu Thr Gly Asp Ser Asn Ala Ile Ser Trp Trp Asp Gly Leu Ile Gly
 51 50 55 60
 53 cct ggc aaa gca ctc gac acc aac cgc tac tgc atc cta tgc acc aac 240
 54 Pro Gly Lys Ala Leu Asp Thr Asn Arg Tyr Cys Ile Leu Cys Thr Asn
 55 65 70 75 80
 57 gtg ctc gga gga tgc aaa gga tcc acc gga ccg agc agt cca cac cca 288
 58 Val Leu Gly Gly Cys Lys Gly Ser Thr Gly Pro Ser Ser Pro His Pro
 59 85 90 95

(pg.6)

RAW SEQUENCE LISTING

DATE: 03/08/2005

PATENT APPLICATION: US/10/525,674

TIME: 08:02:24

Input Set : A:\Final Sequence List - 13111-2-US.txt

Output Set: N:\CRF4\03082005\J525674.raw

```

61 gac gga aaa cca tgg gga tcc aga ttt cca gcc ctt tca atc cgt gac 336
62 Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Ala Leu Ser Ile Arg Asp
63 100 105 110
65 ctt gtc aat gcc gaa aaa caa ctt ttc gac cac ctc ggc atc aat aaa 384
66 Leu Val Asn Ala Glu Lys Gln Leu Phe Asp His Leu Gly Ile Asn Lys
67 115 120 125
69 att cac gca atc atc ggc gga tcc atg gga ggc gca cgc acc ctc gaa 432
70 Ile His Ala Ile Ile Gly Gly Ser Met Gly Gly Ala Arg Thr Leu Glu
71 130 135 140
73 tgg gct gca ctc cac cca cac atg atg acg act gga ttc gtc ata gca 480
74 Trp Ala Ala Leu His Pro His Met Met Thr Thr Gly Phe Val Ile Ala
75 145 150 155 160
77 gtc tca gca cgc gca agc gct tgg caa atc ggt att caa act gca caa 528
78 Val Ser Ala Arg Ala Ser Ala Trp Gln Ile Gly Ile Gln Thr Ala Gln
79 165 170 175
81 atc agc gcc ata gaa ctc gac ccc cac tgg aac ggc ggc gat tac tac 576
82 Ile Ser Ala Ile Glu Leu Asp Pro His Trp Asn Gly Gly Asp Tyr Tyr
83 180 185 190
85 agc ggt cac gca cca tgg gaa gga atc gcc gcc gct cgc cgg atc gcc 624
86 Ser Gly His Ala Pro Trp Glu Gly Ile Ala Ala Ala Arg Arg Ile Ala
87 195 200 205
89 cac ctc acc tat cgc ggc gaa cta gaa ata gac gaa cga ttc ggc act 672
90 His Leu Thr Tyr Arg Gly Glu Leu Glu Ile Asp Glu Arg Phe Gly Thr
91 210 215 220
93 tcc gca caa cac ggt gaa aac cca ctc ggc ccc ttc cga gat cca cat 720
94 Ser Ala Gln His Gly Glu Asn Pro Leu Gly Pro Phe Arg Asp Pro His
95 225 230 235 240
97 caa cgt ttt gcg gtc acg agc tac ctc caa cac caa ggc atc aaa ctc 768
98 Gln Arg Phe Ala Val Thr Ser Tyr Leu Gln His Gln Gly Ile Lys Leu
99 245 250 255
101 gct caa cga ttc gat gca ggt agt tac gtc gtg ctt acc gaa gcc ctc 816
102 Ala Gln Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Ala Leu
103 260 265 270
105 aat cgt cat gac atc gga cgc ggc cga ggc gga ctc aac aaa gcc ctc 864
106 Asn Arg His Asp Ile Gly Arg Gly Arg Gly Gly Leu Asn Lys Ala Leu
107 275 280 285
109 agc gca atc aca gtc ccc atc atg att gct ggc gtt gat acc gat att 912
110 Ser Ala Ile Thr Val Pro Ile Met Ile Ala Gly Val Asp Thr Asp Ile
111 290 295 300
113 ctc tac ccc tat cac cag caa gaa cac cta tca cga aat cta ggc aac 960
114 Leu Tyr Pro Tyr His Gln Gln Glu His Leu Ser Arg Asn Leu Gly Asn
115 305 310 315 320
117 cta ctc gct atg gca aaa atc agc tca cca gta ggc cac gac gct ttc 1008
118 Leu Leu Ala Met Ala Lys Ile Ser Ser Pro Val Gly His Asp Ala Phe
119 325 330 335
121 ctc aca gaa ttc cga caa atg gag cga atc cta aga cat ttc atg gag 1056
122 Leu Thr Glu Phe Arg Gln Met Glu Arg Ile Leu Arg His Phe Met Glu
123 340 345 350
125 ctt tcg gaa gga atc gac gat tcc ttc cga acc aaa cta gag cgc 1101

```

RAW SEQUENCE LISTING

DATE: 03/08/2005

PATENT APPLICATION: US/10/525,674

TIME: 08:02:24

Input Set : A:\Final Sequence List - 13111-2-US.txt

Output Set: N:\CRF4\03082005\J525674.raw

```

126 Leu Ser Glu Gly Ile Asp Asp Ser Phe Arg Thr Lys Leu Glu Arg
127          355          360          365
129 tga
132 <210> SEQ ID NO: 2
133 <211> LENGTH: 367
134 <212> TYPE: PRT
135 <213> ORGANISM: Corynebacterium diphtheriae
137 <400> SEQUENCE: 2
138 Met Leu Thr Thr Thr Gly Thr Leu Thr His Gln Lys Ile Gly Asp Phe
139   1          5          10          15
141 Tyr Thr Glu Ala Gly Ala Thr Leu His Asp Val Thr Ile Ala Tyr Gln
142          20          25          30
144 Ala Trp Gly His Tyr Thr Gly Thr Asn Leu Ile Val Leu Glu His Ala
145          35          40          45
147 Leu Thr Gly Asp Ser Asn Ala Ile Ser Trp Trp Asp Gly Leu Ile Gly
148          50          55          60
150 Pro Gly Lys Ala Leu Asp Thr Asn Arg Tyr Cys Ile Leu Cys Thr Asn
151   65          70          75          80
153 Val Leu Gly Gly Cys Lys Gly Ser Thr Gly Pro Ser Ser Pro His Pro
154          85          90          95
156 Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Ala Leu Ser Ile Arg Asp
157          100          105          110
159 Leu Val Asn Ala Glu Lys Gln Leu Phe Asp His Leu Gly Ile Asn Lys
160          115          120          125
162 Ile His Ala Ile Ile Gly Gly Ser Met Gly Gly Ala Arg Thr Leu Glu
163          130          135          140
165 Trp Ala Ala Leu His Pro His Met Met Thr Thr Gly Phe Val Ile Ala
166 145          150          155          160
168 Val Ser Ala Arg Ala Ser Ala Trp Gln Ile Gly Ile Gln Thr Ala Gln
169          165          170          175
171 Ile Ser Ala Ile Glu Leu Asp Pro His Trp Asn Gly Gly Asp Tyr Tyr
172          180          185          190
174 Ser Gly His Ala Pro Trp Glu Gly Ile Ala Ala Ala Arg Arg Ile Ala
175          195          200          205
177 His Leu Thr Tyr Arg Gly Glu Leu Glu Ile Asp Glu Arg Phe Gly Thr
178          210          215          220
180 Ser Ala Gln His Gly Glu Asn Pro Leu Gly Pro Phe Arg Asp Pro His
181 225          230          235          240
183 Gln Arg Phe Ala Val Thr Ser Tyr Leu Gln His Gln Gly Ile Lys Leu
184          245          250          255
186 Ala Gln Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Ala Leu
187          260          265          270
189 Asn Arg His Asp Ile Gly Arg Gly Arg Gly Gly Leu Asn Lys Ala Leu
190          275          280          285
192 Ser Ala Ile Thr Val Pro Ile Met Ile Ala Gly Val Asp Thr Asp Ile
193          290          295          300
195 Leu Tyr Pro Tyr His Gln Gln Glu His Leu Ser Arg Asn Leu Gly Asn
196 305          310          315          320
199 Leu Leu Ala Met Ala Lys Ile Ser Ser Pro Val Gly His Asp Ala Phe

```

RAW SEQUENCE LISTING

DATE: 03/08/2005

PATENT APPLICATION: US/10/525,674

TIME: 08:02:24

Input Set : A:\Final Sequence List - 13111-2-US.txt

Output Set: N:\CRF4\03082005\J525674.raw

```

200          325          330          335
202 Leu Thr Glu Phe Arg Gln Met Glu Arg Ile Leu Arg His Phe Met Glu
203          340          345          350
205 Leu Ser Glu Gly Ile Asp Asp Ser Phe Arg Thr Lys Leu Glu Arg
206          355          360          365
209 <210> SEQ ID NO: 3
210 <211> LENGTH: 1149
211 <212> TYPE: DNA
212 <213> ORGANISM: Mycobacterium leprae
214 <220> FEATURE:
215 <221> NAME/KEY: CDS
216 <222> LOCATION: (1)..(1146)
217 <223> OTHER INFORMATION: RML02951
219 <220> FEATURE:
220 <221> NAME/KEY: unsure
222 <222> LOCATION: 224 .. 224
223 <223> OTHER INFORMATION: All occurrences of n indicate any nucleotide
225 <400> SEQUENCE: 3
226 atg aca atc tcc aag gtc cct acc cag aag ctg ccg gcc gaa ggc gag 48
227 Met Thr Ile Ser Lys Val Pro Thr Gln Lys Leu Pro Ala Glu Gly Glu
228 1 5 10 15
230 gtc ggc ttg gtc gac atc ggc tca ctt acc acc gaa agc ggt gcc gtc 96
231 Val Gly Leu Val Asp Ile Gly Ser Leu Thr Thr Glu Ser Gly Ala Val
232 20 25 30
234 atc gac gat gtc tgc atc gcc gtt cag cgc tgg ggg gaa ttg tgc ccc 144
235 Ile Asp Asp Val Cys Ile Ala Val Gln Arg Trp Gly Glu Leu Ser Pro
236 35 40 45
238 acg cga gac aac gta gtg atg gta ctg cat gca ctc acc ggt gac tgc 192
239 Thr Arg Asp Asn Val Val Met Val Leu His Ala Leu Thr Gly Asp Ser
240 50 55 60
W--> 242 cac atc acc ggg ccc gcc gga ccg gga cat cnc aca ccc ggc tgg tgg 240
243 His Ile Thr Gly Pro Ala Gly Pro Gly His Xaa Thr Pro Gly Trp Trp
244 65 70 75 80
246 gac tgg ata gct gga ccg ggt gca cca atc gac acc aac cgc tgg tgc 288
247 Asp Trp Ile Ala Gly Pro Gly Ala Pro Ile Asp Thr Asn Arg Trp Cys
248 85 90 95
250 gcg ata gcc acc aac gtg ctg ggc ggt tgc cgt ggc tcc acc ggc cct 336
251 Ala Ile Ala Thr Asn Val Leu Gly Gly Cys Arg Gly Ser Thr Gly Pro
252 100 105 110
254 agt tgc ctt gcc cgc gac gga aag cct tgg ggt tca aga ttt ccg ctg 384
255 Ser Ser Leu Ala Arg Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Leu
256 115 120 125
258 ata tct ata cgc gac cag gta gag gca gat atc gct gca ctg gcc gcc 432
259 Ile Ser Ile Arg Asp Gln Val Glu Ala Asp Ile Ala Ala Leu Ala Ala
260 130 135 140
262 atg gga att aca aag gtt gcc gcc gtc gtt gga gga tct atg ggc ggg 480
263 Met Gly Ile Thr Lys Val Ala Ala Val Val Gly Gly Ser Met Gly Gly
264 145 150 155 160
266 gcg cgt gca ctg gaa tgg atc atc ggc cac ccg gac caa gtc ccg gcc 528

```

RAW SEQUENCE LISTING

DATE: 03/08/2005

PATENT APPLICATION: US/10/525,674

TIME: 08:02:24

Input Set : A:\Final Sequence List - 13111-2-US.txt

Output Set: N:\CRF4\03082005\J525674.raw

```

267 Ala Arg Ala Leu Glu Trp Ile Ile Gly His Pro Asp Gln Val Arg Ala
268           165           170           175
270 ggg ctg ttg ctg gcg gtc ggt gtg cgc gcc acc gcc gac cag atc ggc 576
271 Gly Leu Leu Leu Ala Val Gly Val Arg Ala Thr Ala Asp Gln Ile Gly
272           180           185           190
274 acc caa acc acc caa atc gca gcc atc aag aca gac ccg aac tgg caa 624
275 Thr Gln Thr Thr Gln Ile Ala Ala Ile Lys Thr Asp Pro Asn Trp Gln
276           195           200           205
278 ggc ggt gac tac tac gag aca ggg agg gca cca gag aac ggc ttg aca 672
279 Gly Gly Asp Tyr Tyr Glu Thr Gly Arg Ala Pro Glu Asn Gly Leu Thr
280           210           215           220
282 att gcc cgc cgc ttc gcc cac ctg acc tac cgc agc gag gtc gag ctc 720
283 Ile Ala Arg Arg Phe Ala His Leu Thr Tyr Arg Ser Glu Val Glu Leu
284 225           230           235           240
286 gac acc cgg ttt gcc aac aac aac caa ggc aat gag gac ccg gcg acg 768
287 Asp Thr Arg Phe Ala Asn Asn Asn Gln Gly Asn Glu Asp Pro Ala Thr
288           245           250           255
290 ggc ggg cgt tac gca gtg cag agt tac cta gag cac cag ggt gac aag 816
291 Gly Gly Arg Tyr Ala Val Gln Ser Tyr Leu Glu His Gln Gly Asp Lys
292           260           265           270
294 cta ttg gcc cgc ttt gac gca ggc agc tac gtg gtc ttg acc gaa acg 864
295 Leu Leu Ala Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Thr
296           275           280           285
298 ctg aac agc cac gac gtt ggc cgg ggc cgc gga ggg atc ggt aca gcg 912
300 Leu Asn Ser His Asp Val Gly Arg Gly Arg Gly Gly Ile Gly Thr Ala
301           290           295           300
303 ctg cgc ggg tgc ccg gta ccg gtg gtg gtg ggt ggc att acc tcg gat 960
304 Leu Arg Gly Cys Pro Val Pro Val Val Val Gly Gly Ile Thr Ser Asp
305 305           310           315           320
307 cgg ctc tac cca ctg cgc ttg cag cag gag ctg gcc gag atg ctg ccg 1008
308 Arg Leu Tyr Pro Leu Arg Leu Gln Gln Glu Leu Ala Glu Met Leu Pro
309           325           330           335
311 ggc tgc acc ggg ctg cag gtt gta gac tcc acc tac ggg cac gac ggc 1056
312 Gly Cys Thr Gly Leu Gln Val Val Asp Ser Thr Tyr Gly His Asp Gly
313           340           345           350
315 ttc ctg gtg gaa tcc gag gcc gtc ggc aaa ttg atc cgt caa acc ctc 1104
316 Phe Leu Val Glu Ser Glu Ala Val Gly Lys Leu Ile Arg Gln Thr Leu
317           355           360           365
319 gaa ttg gcc gac gtg ggt tcc aag gaa gac gcg tgt tcg caa 1146
320 Glu Leu Ala Asp Val Gly Ser Lys Glu Asp Ala Cys Ser Gln
321           370           375           380
323 tga 1149
326 <210> SEQ ID NO: 4
327 <211> LENGTH: 382
328 <212> TYPE: PRT
329 <213> ORGANISM: Mycobacterium leprae
331 <220> FEATURE:
332 <221> NAME/KEY: unsure
333 <222> LOCATION: 75 .. 75

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/525,674

DATE: 03/08/2005
TIME: 08:02:25

Input Set : A:\Final Sequence List - 13111-2-US.txt
Output Set: N:\CRF4\03082005\J525674.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 224 ✓

Seq#:3; Xaa Pos. 75

Seq#:4; Xaa Pos. 75

Seq#:43; Xaa Pos. 13,18,45,59,89,137,145,206,297,320,326,366,384

Seq#:44; Xaa Pos. 13,18,45,59,89,137,145,206,297,320,326,366,384

VERIFICATION SUMMARY

DATE: 03/08/2005

PATENT APPLICATION: US/10/525,674

TIME: 08:02:25

Input Set : A:\Final Sequence List - 13111-2-US.txt

Output Set: N:\CRF4\03082005\J525674.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:242 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:192
M:341 Repeated in SeqNo=3
L:349 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:64
L:4125 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:48
M:341 Repeated in SeqNo=43
L:4294 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0
M:341 Repeated in SeqNo=44